

Personalized Maternal Sleep Quality Assessment: An Objective IoT-based Longitudinal Study

IMAN AZIMI¹, (Student Member, IEEE), OLUGBENGA OTI¹, SINA LABBAF²,
HANNAKAISA NIELA-VILÉN³, ANNA AXELIN³, NIKIL DUTT², (Fellow, IEEE),
PASI LILJEBERG¹, (Member, IEEE), and AMIR M. RAHMANI^{2,4}, (Senior Member, IEEE)

¹Department of Future Technologies, University of Turku, Turku, Finland

²Department of Computer Science, University of California, Irvine, CA, USA

³Department of Nursing Science, University of Turku, Turku, Finland

⁴School of Nursing, University of California, Irvine, CA, USA

Corresponding author: Iman Azimi (e-mail: imaazi@utu.fi).

ABSTRACT Sleep is a composite of physiological and behavioral processes that undergo substantial changes during and after pregnancy. These changes might lead to sleep disorders and adverse pregnancy outcomes. Several studies have investigated this issue; however, they were restricted to subjective measurements or short-term actigraphy methods. This is insufficient for a longitudinal maternal sleep quality evaluation. A longitudinal study *a)* requires a long-term data collection approach to acquire data from everyday routines of mothers and *b)* demands a sleep quality assessment method exploiting a large volume of multivariate data to assess sleep adaptations and overall sleep quality. In this paper, we present an Internet-of-Things based long-term monitoring system to perform an objective sleep quality assessment. We conduct longitudinal monitoring where 20 pregnant mothers are remotely monitored for six months of pregnancy and one month postpartum. To evaluate sleep quality adaptations, we *a)* extract several sleep attributes and study their variations during the monitoring and *b)* propose a semi-supervised machine learning approach to create a personalized sleep model for each subject. The model provides an abnormality score which allows an explicit representation of the sleep quality in a clinical routine, reflecting possible sleep quality degradation with respect to her own data. Sleep data of 13 participants (out of 20) are included in our analysis, as their data are adequate for the study, including 172.15 ± 33.29 days of sleep data per person. Our fine-grained objective measurements indicate the sleep duration and sleep efficiency are deteriorated in pregnancy and notably in postpartum. In comparison to the mid of the second trimester, the sleep model indicates the increase of sleep abnormality at the end of pregnancy (2.87 times) and postpartum (5.62 times). We also show the model enables individualized and effective care for sleep disturbances during pregnancy, as compared to a baseline method.

INDEX TERMS Anomaly Detection, Internet of Things, Longitudinal Study, Maternity Care, Sleep Monitoring, Sleep Quality Assessment

I. INTRODUCTION

SEVERAL physical, physiological, and hormonal adaptations occur during pregnancy to accommodate the developing fetus and to prepare the mother for the delivery [1], [2]. Such variations in the maternal body alter sleep patterns of pregnant women in many ways. In this regard, sleep disturbances are particularly prevalent throughout the pregnancy, including various disorders to maintaining sleep (e.g., insomnia), sleep deprivation, and restless legs syndrome [3]–

[6]. Moreover, sleep patterns of pregnant women might be altered in postpartum months, as they experience new life situations after labor [7].

Studies show that sleep disturbances negatively impact maternal and child health during and after pregnancy [8]. Sleep problems are associated with a high likelihood of poor obstetric outcomes and different diseases such as gestational diabetes, preeclampsia, and stress overload [9]–[11]. Also, they lead to increased risk of preterm birth, intrauterine

growth restriction, and unplanned Caesarean deliveries [12], [13]. Moreover, different studies discussed the correlation between sleep disturbances and postpartum diseases and complications such as depression and damage to the mother-infant relationship [12], [14]. Thus, screening, monitoring, and assessment of maternal sleep quality are essential during pregnancy to alleviate sleep disturbances and prevent its potential complications [12], [15], [16].

Sleep quality is a complex concept that is traditionally evaluated via qualitative attributes (i.e., subjective measurements) and more recently via quantitative attributes (i.e., objective measurements) [17]. Subjective techniques determine perceived sleep quality by inquiring the individuals about their sleep experiences such as sleep duration and disturbances. These techniques are often performed via self-report questionnaires such as the Pittsburgh Sleep Quality Index (PSQI) [18] and Berlin Questionnaire [19]. Those are widely used in sleep quality evaluation of different groups of people as they are relatively straightforward and easy to implement for longitudinal studies. Similar subjective techniques have also been utilized for pregnant women to reveal the impact of pregnancy on maternal sleep [5], [8], [15], [16], [20]–[22]. However, such subjective methods can be inaccurate and poorly reflect sleep quality level, as the data collection is mostly limited to scheduled interviews, Internet-based surveys, or self-report questionnaires. The shortcomings and poor performance of such methods have been widely discussed in several studies investigating the validity of the subjective sleep quality assessment methods [17], [23]–[25].

Alternatively, objective techniques measure the user's physical and health conditions and translate the results into sleep attributes such as sleep efficiency and sleep stages for further assessment. Polysomnography (PSG) is a conventional test in this regard, where several bio-signals are acquired for sleep analysis [26], [27]. The PSG, as the gold standard of the sleep assessment, has been exploited for sleep disturbances monitoring in pregnancy [4]. However, it is bounded to one or a limited number of nights due to its data acquisition limits. Actigraphy is another objective method that examines sleep quality by monitoring human rest/activity cycles [28]. Data acquisition in actigraphy is more convenient and non-invasive for users, as it is performed via a small and light-weight wearable device placed on the user's wrist or ankle. Standalone (i.e., without network connectivity and real-time remote access) actigraphy monitors have been deployed for offline and short-time sleep monitoring, such as the works presented in [29]–[32] where maternal sleep is monitored for up to 14 days. However, the constraints in local storage and processing have hindered the utilization of this technology for longitudinal sleep quality monitoring.

Longitudinal objective sleep monitoring necessitates a long-term data collection to acquire data from everyday routines of participants 24/7. We believe recent advancements in Internet-of-Things (IoT) technologies provide an unprecedented opportunity to enable such continuous health

monitoring. IoT is an emerging network of interrelated objects that tailors a distinct set of paradigms such as wearable electronics, communication infrastructure, and data analytics to deliver personalized services to the end-users [33], [34]. However, it should be noted that an IoT-based sleep monitoring system, despite being a powerful tool, generates a large volume of multivariate data which dramatically increases over time. Such big data [35], while being a rich source of information, call for tailored and intelligent data analytic techniques and models.

Conventional techniques assess the sleep quality only from a single perspective by separately extracting and analyzing each sleep attribute (e.g., sleep duration) from a pool of sleep-related data. Data in a high-dimensional space require a more intelligent amalgamation method to transform all sleep attributes into a single overall sleep quality score, in a way that the contribution of each attribute is automatically considered in the final score. This allows a straightforward representation of the sleep quality in a clinical routine and reflecting possible sleep quality degradation of an individual with respect to her own life situation and health condition. We believe that such a method is particularly essential for maternal sleep quality assessment and individualized care approach, as a mother's physical and mental states undergo a process of change throughout the course of pregnancy and postpartum, which necessitates an explicit indicator of the mother's sleep changes during this period.

In this paper, we present an IoT-based long-term monitoring system that employs a wrist-worn device to assess the sleep of pregnant women during pregnancy and postpartum thoroughly. Our monitoring system is deployed on a real human subject trial where 20 pregnant women are remotely and continuously monitored for six months of pregnancy and one month postpartum. We first study sleep quality changes in this monitoring, leveraging several objective attributes. We then propose an anomaly detection approach to construct a personalized sleep model for each individual using the sleep data from the beginning of the monitoring process. We measure the sleep adaptations of the rest of the pregnancy and postpartum, using the personalized model to investigate the maternal sleep quality from a different perspective. In summary, the contribution of this paper is manifold:

- i) Presenting an IoT-based long-term monitoring system to perform objective sleep quality assessment during pregnancy and postpartum.
- ii) Conducting a longitudinal study on a human subject trial on maternal sleep.
- iii) Observing the degradation of sleep quality during pregnancy and postpartum separately for a set of fine-grained quantitative sleep attributes.
- iv) Proposing a neural network-based approach to investigate maternal sleep quality adaptations in a comprehensive and personalized way.

The rest of the paper is organized as follows. We outline the background and related work of this research in Sec-

tion II. Section III describes the study design. In Section IV, we present our sleep analysis approaches. Results and findings are presented in Section V. In Section VI, we discuss our findings, evaluate the model, and represent limitations and future directions of this study. Finally, Section VII concludes the paper.

II. BACKGROUND AND RELATED WORK

In this section, we first outline the background of maternal health and sleep monitoring. Then, we present state-of-the-art anomaly detection techniques as appropriate tools to create models for abnormality detection.

A. MATERNAL HEALTH AND SLEEP MONITORING

Maternal health can be monitored during pregnancy to ensure the well-being of both the mother and her future child. Pregnancy is a window to a woman's future health [36], and thus women are also interested in monitoring their health during pregnancy. Furthermore, using systematic and regular monitoring, several abnormalities and complications regarding pregnancy could be detected early and be treated accordingly. Maternal health monitoring, however, varies in different countries, and only half of women receive the recommended amount of care during their pregnancy [37]. Therefore, there is a need to develop new solutions that can widen the availability of maternal health monitoring for all pregnant women.

Sleep as an important part of overall maternal health requires particular attention. Multiple hormonal and physiological changes during pregnancy might contribute to sleep problems. For example, nausea, vomiting, or anxiety might cause sleep disturbances in the first trimester of pregnancy. As pregnancy progresses, the frequency and duration of sleep disturbances increase. Frequent urination, backache, leg cramps, and anxiety about delivery are common reasons for compromised sleep in the third trimester.

Sleep disturbances are common during pregnancy and are the risk factors of adverse pregnancy outcomes such as prenatal depression, gestational diabetes, and preterm birth [11], [16], [38], [39]. Also, many women suffer from acute sleep deprivation during the postpartum period, and compromised sleep may continue even several months after birth [39]. This problem might lead to diseases such as maternal fatigue and postpartum depression [14]. It is possible to use nonpharmacological strategies such as regular physical activity, controlling weight gain, and relaxation, to alleviate sleep disorders during pregnancy. Medication should be used only in severe cases to avoid possible teratogenic effects [40]. Sleep quality assessment is the first step for managing sleep disturbances and disorders. It gives an accurate picture of sleep changes and assists to early-detect sleep problems [41]. In particular, systematic and personalized sleep assessment enables the provision of right strategies to manage sleep disturbances and disorders of each woman.

Different methods have been proposed in the literature to investigate sleep problems. The duration, as well as the

quality of sleep during pregnancy, has usually been measured using questionnaires [16], [42], [43]. The Pittsburgh Sleep Quality Index (PSQI) is the gold standard for subjective sleep quality assessment, in which individuals are asked to answer a self-report questionnaire [18]. The tool discriminates "good" sleep quality from "bad" leveraging seven component scores such as sleep latency, habitual sleep efficiency, and use of sleeping medication. Such subjective methods are not accurate; pregnant women have both over and underestimated their sleep duration compared with objective measurements [44].

Polysomnography (PSG) is the gold standard of sleep monitoring. The method typically employs various wearable sensors to capture several bio-signals including electroencephalogram (EEG), electromyogram (EMG), electrooculogram (EOG), and electrocardiogram (ECG), providing different sleep indices such as sleep efficiency, sleep onset latency, and sleep stages [4], [26], [45]. However, the use of the PSG is limited to sleep laboratories and clinical settings due to the burdensome implementation of its multisensor data acquisition. Therefore, the method was mostly performed in a short period of time in sleep studies. For example, an overnight lab-based PSG was implemented along with the Berlin questionnaire, targeting obstructive sleep apnea [19]. Similarly, in the maternity care, sleep disturbance was investigated via a short-term PSG-based data collection, i.e., two consecutive nights in each trimester, and in first and third postpartum months [4].

Actigraphy is another low-cost alternative for monitoring sleep and sleep-wake behavior of an individual [28]. The Sleep actigraphy typically includes an actigraph device equipped with a 3-axis MEMS accelerometer sensor, a low-performance processor and a limited memory. The acceleration data are locally processed, and sleep parameters are extracted. The actigraphy method is easy-to-use in out-of-hospital settings in contrast to the PSG. However, it is bounded to offline services. Objective sleep monitoring has been fulfilled in different maternal studies using short-term actigraphy methods [30], [46]. For example, Lee and Gay [29] investigated the association between sleep disturbance in late pregnancy with labor using an actigraphy for 2 days along with subjective measurements in the ninth month of pregnancy; a seven-day actigraphy and PSQI methods were employed for maternal sleep disturbance [31]; and Haney *et al.* [32] assess sleep in early pregnancy exploiting a 14-day actigraphy method, questionnaires, and blood pressure measurements.

Contact-free sensors have also been proposed for sleep monitoring. Some examples are visual-based sensors [47], mattress-based sensors [48], and smartphone sensors [49]. They were mostly designed to acquire sleep patterns as well as vital signs such as heart rate and respiration rate. The use of such systems has been limited in real-world applications because of restrictions in data collection and high cost. In one study, the maternal body movements of 2 pregnant women were monitored for a couple of weeks, using a piezoelectric

sensor board placed beneath their mattress [50].

B. ANOMALY DETECTION

Anomaly detection, also known as outlier detection, is the problem of finding patterns or events in data that differ from the expected behavior [51]. Anomaly detection has been applied in many fields including fraud detection, healthcare, and intrusion detection in cybersecurity [52]. An anomaly detection technique applied to a problem depends on a variety of factors including the availability of labeled data, the nature of the data, the type of anomalies to be detected, the output of the method, and in some cases the field of study.

The type of anomalies in a dataset can be divided into three major categories [51]. First, *point anomalies* refer to data instances that are anomalous with respect to the rest of the data (i.e., normal data). Second, *contextual anomalies* are data instances that are anomalous in a certain context. For example, 150 heart beats per minute would be normal during exercise although it is anomalous if the user is sleeping. Third, *collective anomalies* refer to a group of related data instances which together are considered anomalous. For instance, recording a couple of high heart rate events in a day would be detected as anomalous (e.g., health deterioration) in a health application. Moreover, datasets can be modified to change the anomaly type; e.g., *point anomalies* and *collective anomalies* can become *contextual anomalies* if we add context information to the dataset.

The choice of a specific anomaly detection method – *supervised*, *semi-supervised*, and *unsupervised* – is greatly dependent on the type of data involved. The data can generally be divided into binary, categorical, or continuous. However, it can be a combination of these categories in some cases. In addition, the output of the method can be either binary (i.e., normal or anomalous) or continuous in the form of an anomaly score which represents the degree of the anomaly [51]. The availability of labeled data is a common challenge in anomaly detection, as anomalies might not occur frequently. Moreover, labeling of a dataset by an expert is time-consuming and expensive. The extent of the availability of a labeled dataset determines which method is used.

Supervised anomaly detection methods rely on data with labels for both the normal and anomalous classes. They construct a predictive model to differentiate normal and abnormal behavior. However, unbalanced distribution of data should be considered in such models, as in practice anomalous data do not occur as often as normal data. Examples of such methods include Neural Networks methods [53], Support Vector Machine (SVM) [54], and Rule-based approaches (e.g. Decision trees) [55].

Semi-supervised anomaly detection methods deploy semi-supervised learning (also known as one-class learning methods) that only consider normal data to train their models. When the model is created to understand normal behavior, it can then distinguish between normal and anomalous classes. These methods are commonly applied because of unavailability or shortage of anomalous data in many applications.

Moreover, no data labeling is required, as all the input data are normal. Some examples of these methods are Statistical techniques [56], one-class Support Vector Machine (SVM) [57], and Neural Networks methods [58]–[60].

In contrast, *unsupervised* anomaly detection methods deploy unsupervised learning techniques that require no training data, assuming the normal data occur more often than anomalous data. Unfortunately, applying data that do not fit this assumption would lead to a high false positive rate. Clustering techniques [61] and Nearest Neighbor techniques [62] are examples of unsupervised or semi-supervised techniques, which rely on the assumption that normal data remain in a cluster or dense neighborhood while anomalous data do not. They often require large training data for the normal classes.

III. STUDY DESIGN

This paper proposes an IoT-based monitoring system equipped with a semi-supervised machine learning approach, by which pregnant women can be monitored remotely, continuously, and long-term. Also, the proposed system enables personalized sleep analysis during pregnancy and the postpartum, providing effective care for maternal sleep disturbances. We present this system for a real human subject trial on maternal sleep, where pregnant women are monitored in six months of pregnancy and one month postpartum. In this section, we introduce the IoT-based monitoring system and provide details about our implementation setup, the participants, and recruitment.

A. IOT-BASED MONITORING SYSTEM

An IoT-based system is introduced to continuously monitor the pregnant women. As shown in Figure 1, the architecture of the proposed system is partitioned into three main tiers. First, the sensor network performs data collection in IoT-based systems, located in the vicinity of the end-users. It acquires pregnancy- and sleep- related data from the end-users constantly. Thanks to the advances in embedded and wearable technologies, various lightweight energy-efficient wearable devices such as smartwatches, fitness trackers and Holter monitors are nowadays available for this tier.

The gateway, as the second tier, is a bridge between the sensor network and the Internet (i.e., cloud servers). The gateway is responsible for data transmission and protocol conversion. Smartphones and tablets as widespread mobile computing devices can be employed in this layer. They provide data transmission in both directions, transmitting collected health data to the cloud servers as well as sending reports and feedback to the end-user. Moreover, subjective measurements including interviews and Internet-based surveys can be carried out.

The cloud server, as the third tier, includes a high-performance computing infrastructure. It is responsible for the sleep quality analysis (e.g., data abstraction and modeling). Our semi-supervised machine learning approach is fully positioned at this tier. Moreover, the cloud server manages, secures, and stores the data remotely and is capable of pro-

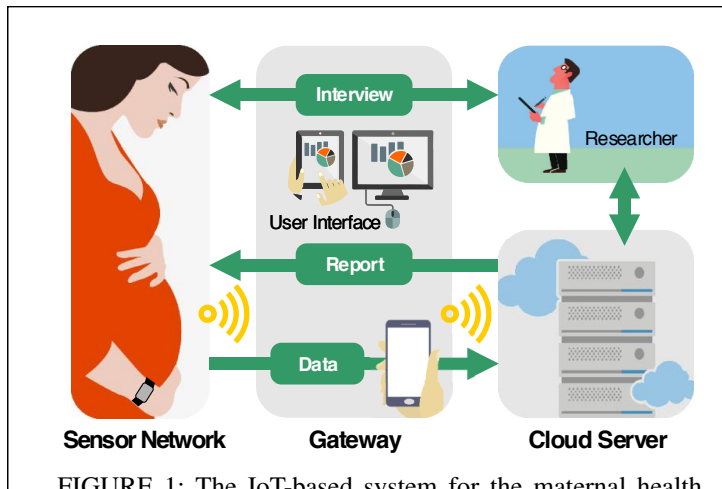


FIGURE 1: The IoT-based system for the maternal health monitoring.

viding a control panel for data visualization. The processed data are shared with the experts (e.g., researchers) for further analysis.

Setup: For the data collection, we restricted our selection of sensor nodes to wearable products (e.g., smart wristbands and smartwatches) that are technically applicable and practically feasible to continuous long-term monitoring [63]. Various studies have shown the validity and reliability of such wearables in terms of sleep parameters by comparing different wearables with the gold standard PSG [64]–[66]. At the beginning of the study, various devices such as Garmin Vivosmart HR [67], Microsoft Band¹ and Fitbit Charge HR² were available in the market. We selected the Garmin Vivosmart HR considering several factors such as the built-in sensors, battery life, small size, light weight, strap design, and waterproofness. More details of the feasibility of this study can be found in [68].

The Garmin Vivosmart HR contains an optical sensor and an inertial measurement unit (IMU), through which photoplethysmogram (PPG) [69] and acceleration signals are collected. In our setup, the participants were requested to wear the device continuously. We acquired a set of data every 15 minutes, including heart rate, step counts, and body movements. The data were utilized for the sleep analysis.

In addition, the pregnant women were asked to frequently synchronize the wristband's data with the remote servers via gateway devices – their smartphones or personal computers in this setup. For the server, we used a Linode virtual private server (VPS) [70] with two 2.50GHz Intel Xeon CPU (E5-2680 v3), 4GB memory, and SSD storage drive. The cloud server was used to store the data remotely, to perform the sleep quality analysis methods, and to provide data visualization.

TABLE 1: Background information of the selected participants.

Statement	Type	Value
Age at pregnancy (years)	–	25.7 ± 4.9
Pre-pregnancy BMI	–	25.0 ± 6.4
Quantity of pre-pregnancy physical activity in week	Once or less Sometimes Almost daily	3 women 5 women 12 women
Quality of pre-pregnancy physical activity in week	Light Moderate Vigorous	8 women 11 women 1 woman
Employment Status	At work Student Unemployed	13 women 5 women 2 women
Marital Status	Married or with partner Single	17 women 3 women
Educational Status	Below secondary education Secondary education College University	4 women 9 women 4 women 3 women
Smoking Status	Pre-pregnancy In-pregnancy	7 women 5 women

B. PARTICIPANTS AND RECRUITMENT

The monitoring was performed on primiparous pregnant women attending to one of two selected maternity outpatient clinics in Southern Finland Between May 2016 and June 2017. Practically, all pregnant women in Finland visit a public health nurse regularly in a maternal health clinic. They may also participate in a free of charge ultrasound examination at the end of first trimester. The participants of this study were recruited in this examination satisfying certain criteria:

- 1) The participant is at least 18 years old.
- 2) She should expect her first child.
- 3) The pregnancy is singleton.
- 4) The gestational age should be less than 15 weeks.
- 5) She understands Finnish or English.
- 6) She owns a smartphone, tablet, or personal computer.

Twenty-two pregnant women who met the criteria were informed after the ultrasound examination. Based on this initial interest, the procedure and purpose of the study were provided for the women with phone calls. Twenty women agreed to participate in the study. In face-to-face meetings, the researchers collected background information of the participants, some of which presented in Table 1. Afterward, the wearable devices and instructions were delivered to the participants.

C. ETHICS

The study was conducted in accordance with the code of ethics of the World Medical Association (Declaration of Helsinki) for involving human subjects in the experiments. It was also approved by the joint ethics committee of the hospital district of Southwest Finland (35/1801/2016) and Turku University Hospital (TYKS). Moreover, the written

¹<https://support.microsoft.com/en-us/help/4000514/band-2-get-started>

²<https://www.fitbit.com/be/chargehr>

informed consent was obtained from all participants enrolled. In addition, the permission to use Garmin Vivosmart® HR (Garmin Ltd, Schaffhausen, Switzerland) in this study was acquired from the manufacturer Garmin Ltd.

IV. SLEEP QUALITY ANALYSIS

In this section, we present our sleep quality analysis approach tailored for assessment of maternal sleep adaptations during pregnancy and postpartum. From the collected data, we first extract several sleep attributes, each of which focuses on a specific aspect of sleep quality. Changes and trends of these attributes are explored for each subject throughout the monitoring process. We then propose a personalized sleep model for each subject to assess sleep quality in a comprehensive and personalized way. The personalized model is constructed by feeding the sleep attributes from the early stages of the monitoring to a machine learning approach.

A. SLEEP ATTRIBUTES

Various objective sleep attributes have been proposed in the literature for sleep quality assessment at many levels [71]. The selection of these attributes depends on the type of collected data (i.e., bio-signals and acceleration data) and subsequently the level of the analysis. For example, actigraphy can be used to extract sleep quantity parameters such as sleep duration and awake after sleep onset [72], [73]. On the other hand, EEG, EOG, and respiration signals are utilized to obtain attributes related to the sleep stages (e.g., REM sleep) [74]. In this study, a wristband equipped with PPG and IMU sensors is employed to continuously collect different parameters such as physical activity, body movements, and heart rates. We exploit these parameters to extract conventional sleep quantity, quality, and schedule attributes [17], [23], [71], [75]. In this regard, eight objective sleep attributes are extracted from each sleep event during nighttime to investigate maternal sleep adaptations. The attributes are outlined as follows:

- *Sleep Duration*, also known as Total Sleep Time (TST), indicates the total time that a user sleeps in a day [76]. It is one of the prevalent parameters in sleep analysis, widely used as a predictor of illnesses and mortality. The association between short/long sleep duration and high risks of different diseases such as cardiovascular diseases, stroke, and hypertension is demonstrated in the literature [77], [78]. In this study, the sleep duration is extracted using sleep information (i.e., start and end of the sleep) provided by the Garmin Vivosmart HR. To validate the sleep information, we implemented a manual cross-check between the sleep information and other data such as body movements and heart rates. The sleep information is corrected or discarded if there was no match between the data. Note that a Listwise deletion method is used to eliminate sleep events including missing values [79]. We also excluded short naps in the analysis, due to the limitations of our study.
- *Sleep Onset Latency (SOL)* refers to the amount of time that a user spends in bed before her status changes to the sleep state [80]. In this study, the sleep onset latency is obtained using the step counts data, and body movements and orientations. It is the time between the occurrence of the last step before the sleep event and the beginning of the sleep event.
- *Wake After Sleep Onset (WASO)* refers to the amount of time that a user is awake after the sleep has begun and before the final awakening [80]. In this study, we use body movements and orientations data to determine the WASO during the sleep event. Step counts data are also used to detect if the user leaves the bed.
- *Sleep Fragmentation* indicates the number of awakenings that occur after the sleep is initiated and before the final awakening [81]. In this study, the sleep fragmentation is also obtained using the body movements and step counts data, by counting the times the user wakes or leaves the bed during the sleep event.
- *Sleep Efficiency* is the ratio of the time that the user is sleeping (i.e., sleep duration) to the total time spent in bed [4]. In this study, the bedtime is determined using the step counts data. It is considered as the time between the occurrence of the last step before the sleep event and the first step after the sleep event. The sleep efficiency is calculated as sleep duration divided by bedtime.
- *Sleep Depth* reflects the ratio of deep sleep duration (i.e., motionless sleep) to the amount of time of total sleep (i.e., sleep duration). Conventionally, the sleep stages including non-REM (i.e., N1, N2, N3, and N4 stages) and REM sleep are measured via Polysomnography tests utilizing EEG, EMG, and EOG signals [82], [83]. However, due to limitations of the data collection in this long-term monitoring, these sleep stages cannot be distinguished. In this study, this attribute is defined according to the body movements data, showing the amount of motionless sleep in total sleep period, which likely reflects deep sleep (i.e., N3 and N4 stages).
- *Resting Heart Rate* refers to the number of heart beats per minute when the user is at complete rest. As a cardiovascular risk factor, this attribute was investigated in studies, tackling associations between elevated resting heart rate and increased risk of cardiovascular diseases and mortality [84], [85]. In this study, we define this attribute for each sleep period by calculating the minimum value of total sleep heart rates.
- *Heart Rate Recovery* is the time between the start of the sleep and the time when the resting heart rate is reached. This attribute can be considered as a readiness score of the user. In this study, heart rate recovery is obtained using sleep event and resting heart rate information.

B. PERSONALIZED SLEEP MODEL

We propose a personalized sleep model to investigate sleep quality adaptations in pregnancy and postpartum. The model is trained via the user's sleep data at the beginning of the

monitoring. Then, the model is used to evaluate the changes and trends of data from the rest of the monitoring (i.e., test data). The test data instances are affected by the new life conditions of pregnancy; and as the model output, a score is desirable that is indicative of the degree of the sleep abnormality.

The personalized models for sleep can leverage anomaly detection methods for identifying such abnormalities and outliers in a dataset. We delve into state-of-the-art anomaly detection methods and develop a suitable method for maternal sleep quality assessment. As mentioned in Section II-B, there is a broad range of methods for anomaly detection. However, many of them are inappropriate for our study.

In this monitoring scheme, a data instance or sleep event is multivariate (i.e., multiple attributes), and no contextual or behavioral data is included. Therefore, we only focus on *Point Anomalies* approaches where a data instance can be selected as anomalous with respect to the rest of the data instances, but not the context information. Moreover, the proposed technique should create a model using the “normal” data. Therefore, our selection is narrowed down to semi-supervised anomaly detection techniques.

Considering the output produced by the anomaly detection, binary techniques are not applicable in this study because they assign a binary label (i.e., normal or abnormal) to the test instance. Support vector machine-based methods are examples of binary techniques. Also, rule-based techniques generally require training data to contain labels for both normal and anomalous classes [55]. Moreover, Nearest Neighbor techniques (e.g., KNN) use a distance between a test data instance and its nearest neighbors to determine if it is anomalous. However, their performance highly depends on the size of the training data and dimensionality of the features. Clustering techniques are difficult to apply when the training data is small because there is a high tendency for the anomalous class to form a large cluster leading to a high false positive rate [61]. Statistical techniques present alternatives that rely on the assumptions (i.e., statistical models) made about the data generating distribution. They are also inappropriate since the assumptions tend not to hold true in high-dimensional data (like our dataset) and cannot capture interactions between features [51].

In contrast, artificial neural networks have been successfully applied to anomaly detection in various fields [53], [58], [86]. Replicator Neural Networks (RNN), also known as Auto-encoders, are the most commonly used form of neural networks in semi-supervised and unsupervised settings [58], [86], [87]. They are known for their ability to work well with high dimensional datasets and to capture linear and nonlinear interactions in the data. However, these techniques might show poor performance when the training data size is small.

Bayesian networks-based methods tackle this issue, including probability distributions in their models. They provide an uncertainty estimate along with the output, where it serves as a confidence bound on the output of the model. In addition, the model performs efficiently in case of small

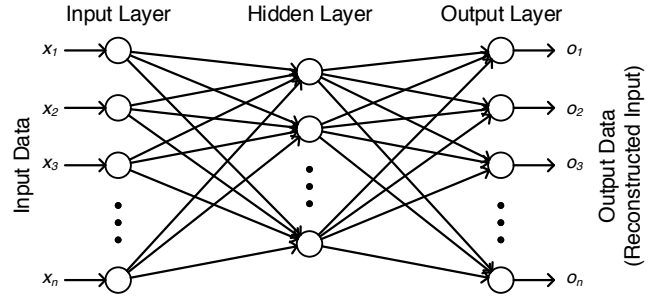


FIGURE 2: Replicator Neural Network with one hidden layer.

data instances and is robust to over-fitting [88]. This quality is important in this study, as we have a limited amount of data samples (i.e., sleep events for each participant) to train an individualized sleep model. Integrating a Bayesian method into artificial neural networks was first proposed by MacKay [89] and Neal [90]. This technique has been applied in several domains including medical diagnostics and Internet traffic classification [91].

We exploit the same concept to construct the personalized sleep model, incorporating a Bayesian approach into a Replicator Neural Networks (RNN).

RNN was first proposed by Hawkins *et al.* [59] and has been further developed by Dau *et al.* [60]. The method belongs to the class of auto-associative Neural Networks with compressed internal representations [60]. It captures a nonlinear representation of the input data and attempts to reproduce the input data as the output of the network. During the training process, the weights in the network are optimized to minimize reconstruction errors of the training data. For a given data instance (i), the reconstruction error is defined as:

$$\delta_i = \frac{1}{n} \sum_{j=1}^n (x_{ij} - o_{ij})^2 \quad (1)$$

where n is the number of features in the data instance, x_{ij} is the input data instance, and o_{ij} is the output of the RNN. The reconstruction error, δ_i , can be used as the anomaly score for the given data instance.

Our Bayesian RNN is designed with one hidden layer, as depicted in Figure 2. Given the training inputs as $X = \{x_1, \dots, x_n\}$ and their corresponding outputs as $Y = \{y_1, \dots, y_n\}$, we aim to find a function, $f^w(X)$ parameterized by weights w , that is likely to generate the outputs. $f^w(x)$ is defined as $f^w(X) = g(W_2 h(X))$, where $h(X)$ is the hidden layer which is $h(X) = g(W_1 X)$. W_1 and W_2 are weights vectors defined over probability distributions; and the activation function is the rectified linear unit (ReLU) (i.e., $g(z) = \max\{0, z\}$).

It should be noted that Bayesian Neural Networks are based on Bayes theorem, and in general we need to find the posterior distribution of the weights. Therefore, we begin by setting a prior probability distribution on the weights, $p(w)$, with a Gaussian probability distribution. We, then,

obtain the likelihood, $p(Y|X, w)$, by updating our beliefs about the prior, $p(w)$, after seeing the data and deciding which weights are more likely to produce the outputs. The posterior distribution $p(w|X, Y)$ is defined over the space of the weights:

$$p(w|X, Y) = \frac{p(Y|X, w)p(w)}{p(Y|X)} \quad (2)$$

where $p(Y|X)$ is the model evidence. However, the posterior distribution cannot be computed by Equation 2, as the model evidence is intractable for most real life problems [88], [92]. Therefore, an approximation method such as Variational Inference [93] is used to obtain an approximating distribution as:

$$q(w) = p(Y|X, w)p(w) \quad (3)$$

$q(w)$ should be as close as possible to the true posterior distribution $p(w|X, Y)$ in Equation 2. Therefore, the Kullback–Leibler (KL) divergence³ [94] of the two distributions must be minimized:

$$KL(q(w)||p(w|X, Y)) = \int q(w) \log \left(\frac{q(w)}{p(w|X, Y)} \right) dw \quad (4)$$

However, Equation 4 still contains the model evidence, so it is still intractable. This leads to the use of Evidence Lower Bound (ELBO) as an alternative to the KL divergence. The ELBO is the negative of the KL divergence up to a logarithm constant. Therefore, maximizing the ELBO is equivalent to minimizing the KL divergence which in turn lets us to approximate the true posterior distribution:

$$\begin{aligned} ELBO &= \int q(w) \log p(Y|X, w) dw - KL(q(w)||p(w)) \\ &\leq \log p(Y|X) \end{aligned} \quad (5)$$

In our Bayesian RNN, we maximize the objective in Equation 5. More details can be found in [88], [92], [95].

V. EXPERIMENTAL DETAILS AND RESULTS

Twenty pregnant women were recruited to participate in this study. The gestational ages of the subjects were 12 ± 2.1 weeks at the beginning of the monitoring. On average, the subjects were 25.7 years old and had pre-pregnancy body mass index (BMI) of 25, with different lifestyles and background characteristics as shown in Table 1.

We excluded 7 participants from our sleep analysis, as they forgot/refused to use the wristband during sleep, with the result that their data were insufficient for our study. Therefore, in the final analysis, 13 pregnant women were included in our analysis. For these 13 subjects, we extracted valid sleep data for 172.15 ± 33.29 days per person out of the total 216.61 ± 14.34 days of the monitoring (79.5%). The

valid sleep data included 76.08 ± 15.17 days of the second trimester per person, 78.69 ± 12.75 days of the third trimester, and 17.38 ± 10.45 days of 1-month postpartum.

Regular phone-interviews (i.e., once or twice a month) were performed during the study to acquire subjective measurements of their status. According to the self-reports, the subjects mostly had their daily routines (i.e., regular work or study) prior to week 30, and began maternity leaves from weeks 30-34 through the end of our study. In addition, the participants were requested to report if they encounter sleep disturbances. On average, three women reported sleep problems at each interview till week-34, and six women experienced difficulty at sleeping in the final weeks of pregnancy. The complaints were mostly due to back pain, sickness, and visiting the toilet during nights.

In the following, we first present the eight objective sleep attributes measured from the participants during pregnancy and the postpartum; then, we demonstrate the abnormality scores calculated using our proposed approach.

A. SLEEP ATTRIBUTES

As discussed in Section IV-A, eight objective sleep attributes are exploited in this study to investigate the maternal sleep changes from different perspectives. To visualize the collected data, we calculate the weekly average of the sleep attributes, where each week contains valid sleep data for at least 4 days. The weeks with less than 4-days data were excluded (4.7 ± 3.6 weeks per person) to reduce the bias.

The variations in attributes for the 13 participants are illustrated in Figures 3, starting from week 13 to week 40 of pregnancy and week 1 to week 4 of postpartum. The variations are depicted by minimum, first-quartile, median, third-quartile, and maximum values of the attributes in each week. Weeks 39, 40, and 41 were the delivery weeks of 3, 7, and 3 participants, respectively. We excluded the data of week 41 in the figures, since we had the sleep data of only one participant.

Sleep duration, a key parameter in sleep quality assessment, gradually decreased during pregnancy. As indicated in Figure 3a, it was 8 hours and 20 minutes (median value) on the weeks 13-15, then decreased by approximately 10% and 20% in the mid and end of third trimester, respectively. It dropped to 5 hours and 50 minutes (median value) on the first week of postpartum and increased afterward.

On the other hand, the *WASO* dramatically increased (see Figure 3b). This parameter was more than 2-times higher at the third trimester and 3-times higher at the postpartum in comparison to the second trimester. Therefore, the quality of sleep diminished at the last stages of pregnancy, and it even became worse after the labor.

Similarly, *sleep fragmentation* increased, so there were more awakening times at the third trimester and postpartum as illustrated in Figure 3c. The variations of the *sleep efficiency* were in accordance with the previous attributes, where it gradually decreased throughout the pregnancy and was at the lowest after the delivery (see Figure 3d).

³KL divergence, written as $KL(p||q) = \int p(x) \log \left(\log \frac{p(x)}{q(x)} \right) dx$, is a measure of the distance between probability distributions in this case p and q . A known property of the KL divergence is that is always greater or equal to zero

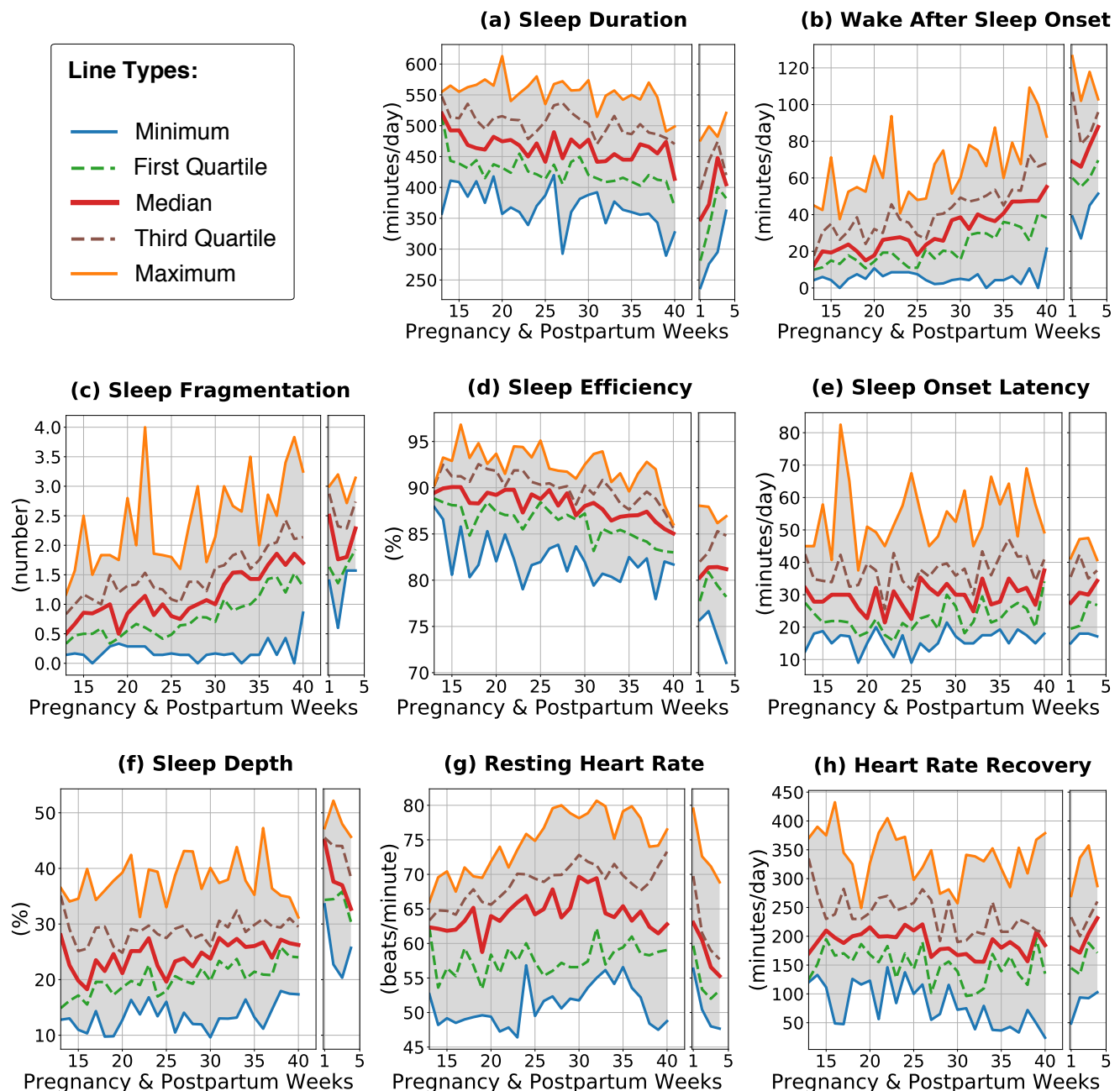


FIGURE 3: The sleep attributes of the 13 participants from week 13 to week 40 of pregnancy and week 1 to week 4 of postpartum. The variations are indicated by minimum, first-quartile, median, third-quartile, and maximum values of the attributes.

The increase in *sleep onset latency* was insignificant during pregnancy. As indicated in Figure 3e, the parameter slightly elevated at the third trimester (on average 30.92 minutes) in comparison to the second trimester (on average 27.69 minutes). In a similar manner, *sleep depth* hardly increased in the pregnancy (see Figure 3f). However, the parameter jumped to more than 40% after the labor. Accordingly, motionless sleep (i.e., deep sleep) was relatively elevated in

postpartum, although the sleep duration was less than sleep duration in the pregnancy period.

The heart-rate-related attributes are depicted in Figures 3g and 3h. *Resting heart rate* increased in the second trimester by more than 10%. However, the parameter was relatively less in postpartum, where it was, on average, 55 beats per minute at the postpartum week 4. As indicated in Figure 3h, *heart rate recovery* also changed during pregnancy. It de-

creased in the third trimester (on average 175.78 minutes) in comparison to the second trimester (on average 201.71 minutes).

B. ABNORMALITY SCORE

Recall that the sleep quality score is computed through an abnormality score using our Bayesian RNN approach. The cloud server is responsible for the sleep model construction (i.e., training phase) and abnormality score calculation (i.e., testing phase). To implement the Bayesian RNN, we use the Lasagne [96] and PyMC3 [97] frameworks in Python. The input data of the method are the sleep data. Each data instance includes the eight sleep attributes of a sleep event during nighttime. The method has one input, one output, and one hidden layers, each of which has eight units (i.e., number of the sleep attributes).

Model Construction: As aforementioned, the training data are the “normal” data in such semi-supervised algorithms. In this study, the user’s sleep data at the beginning of the monitoring were considered as the training data. These are the data from week 13 to week 21, as the most similar data to the user’s normal condition. It should be noted that, in an ideal situation, pre-pregnancy sleep data should be selected as the training dataset (i.e., “normal” data).

The training data were normalized and fed to the model. Using the PyMC3, the weights were first initialized as normal probability distributions and then were optimized by maximizing the Evidence Lower Bound from the Equation 5. Therefore, the model was enabled to replicate the input training data at the output with the minimum error.

Score Calculation: The model, as a compressed representation of the training dataset, was used to reconstruct the test data. In this study, the test data were the sleep data from week 22 to the end of the monitoring. The error of a test instance reconstruction indicates the abnormality level of the test instance. Let us take two different examples. 1) The model replicates the input test data at the output with small error. This indicates the test instance is close to the training dataset (i.e., a similar sample was already seen in the training phase). Consequently, the test instance is “normal”. 2) The model reproduces the input test data at the output with large error. This shows the test instance is far from the training dataset (i.e., the instance is new to the model). Therefore, it is “abnormal”.

In this regard, the abnormality level (i.e., abnormality score) is the distance between the input and reconstructed output, calculated as:

$$s = \frac{1}{n} \sum_{j=1}^n (x_j - o_j)^2 \quad (6)$$

where n is the number of sleep attributes which is 8, x_j is the original data instance and o_j is the reconstructed data instance.

In this work, a personalized RNN model was created for each participant using her own data; and her test data were

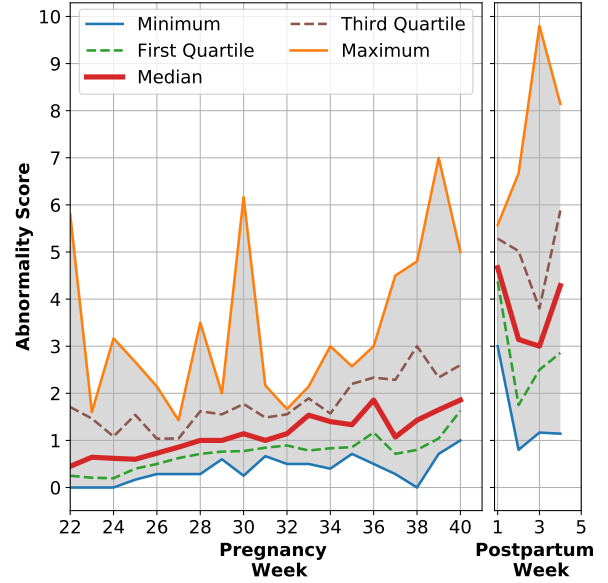


FIGURE 4: The abnormality scores of the 13 participants of pregnancy weeks 22-40 and postpartum weeks 1-4.

evaluated with the personalized model. The *abnormality scores* of the 13 participants are shown in Figure 4, starting from week 22. The overall median values gradually increased as the pregnancy progresses. The highest scores during pregnancy were for week 35 to the labor. At the postpartum week 1, the score jumped to more than 230% in comparison to week 40. This means that the worst sleep quality was for the first week after the labor. Afterward, the scores slightly decreased in the postpartum although they were considerably higher than the scores during the pregnancy.

VI. DISCUSSION AND EVALUATION

To the best of our knowledge, this is the first IoT-based longitudinal study that objectively assesses maternal sleep quality during pregnancy and postpartum. This IoT-based monitoring provides a feasible method to assess the quality of women’s sleep in a challenging transition period from pregnancy to motherhood. In this section, we first discuss the observations made by analyzing each attribute individually and then look into the final sleep abnormality score.

Sleep Attributes: Different objective sleep attributes indicate the quality of sleep diminished during pregnancy and in postpartum. Compared with the existing studies, this work represents a higher *confidence level* on these findings by performing *long-term* and *fine-grained quantitative* measurements and analysis of everyday data of pregnant women.

We found that the sleep duration and sleep efficiency gradually decreased across pregnancy. Correspondingly, the WASO and sleep fragmentation increased. These findings of this continuous wristband monitoring are in concordance with previous knowledge gained from short-term measurements in a few separate time points. Sleep disturbances

during pregnancy could be considered unavoidable due to the hormonal, anatomical, and physiological changes in the woman's body. For example, the levels of oxytocin, prolactin, and cortisol increase and have effects on sleep regulation. Furthermore, respiratory, musculoskeletal, and cardiovascular changes, as well as weight gain and bladder compression by the uterus have impacts on sleep [80].

Moreover, our results indicate there are more changes in these attributes after the delivery. The sleep duration and sleep efficiency drop by 21.5% and 9.7%, and the WASO and sleep fragmentation increase by 3.5 and 4.7 times, in comparison to the second trimester. These postpartum findings also comply with the previous findings; the changed life situation is a common reason for such poor sleep quality. In a previous study by Hughes *et al.* [98], for example, the total sleep time in the first 48 hours after birth was less than 10 hours; however, breastfeeding mothers slept longer than bottle-feeding mothers. Sleep is often compromised in the postpartum period during the first months because of infants' sleep-wake patterns and various needs leading multiple night-time awakenings. Total sleep time appears to be the lowest one month after birth, but it can remain as low still at two months postpartum [39], [99]. In previous studies, these attributes were measured via subjective self-report questionnaires or short-term objective actigraphy [5], [16], [31], [100].

Based on the data in this study, the sleep onset latency did not change significantly during pregnancy; however, the difficulties of falling asleep have been reported to increase as pregnancy progresses [101]. In [101], about one-fourth of pregnant women have suffered from daytime sleepiness which might be an indicator of the insufficient sleep depth. Subjectively rated sleepiness symptoms remained the same during pregnancy [101] as did the sleep depth in this study. Interestingly, the sleep depth increased more than 40% after the delivery. This might be explained with the sleep depth accumulated during pregnancy. Findings related to the heart rate were supported by the earlier knowledge [102]; resting heart rate increased during pregnancy but decreased again during the first month postpartum, and heart rate recovery decreased toward the end of pregnancy.

Abnormality Score: Each sleep attribute represents the maternal sleep quality from a single perspective. We tackled this issue by using an abnormality score which is the fusion of the sleep attributes. It provides a better understanding of changes in maternal individual sleep quality, tailoring sleep data of early pregnancy to evaluate sleep data of late pregnancy and postpartum. In an ideal situation, changes would be evaluated against pre-conception sleep quality [103]. Moreover, it can be used to achieve personalized healthcare. The proposed score enables personalized decision-making through objective sleep quality assessment, where the intensity of the score corresponds to its distance from the user's normal condition (i.e., user's model). This personalization is important in such health-related applications, as the *normal* health condition is

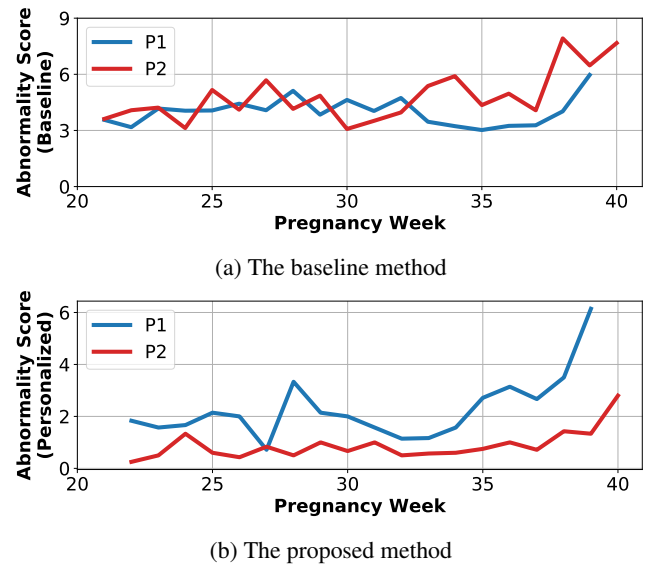


FIGURE 5: The abnormality scores of two participants, using the baseline and proposed methods.

specific for each individual and is not easy to be generally defined. For example, average resting heart rates of two different persons could be 50 and 60 beats/min, both of which are normal values according to their individual conditions.

We evaluate the obtained abnormality scores, comparing the proposed sleep model with a baseline method. Recall that as a semi-supervised approach is used in this work, the training data are labeled as "normal" and the test data are unlabeled. To evaluate the model, we rely on the general hypothesis behind the model, which should produce a higher score in the case of anomalous data (i.e., differentiate "normal" and "abnormal" test instances).

In this regard, we consider a simple aggregate method as a baseline for the performance comparison. The baseline method determines sleep quality scores using overall population values in *normal* conditions. We use the data from the beginning of the monitoring (i.e., *normal* data) representing the most probable sleep attributes of normal conditions in our study. Eventually, the baseline score of each sleep event is the sum of distances between the sleep attributes and their corresponding normal population means in units of the standard deviations.

We select two participants (i.e., P1 and P2) with different conditions to implement the comparison between the proposed method and the baseline. P1 experienced substantial changes in her sleep although P2 had relatively less sleep changes in pregnancy. Table 2 shows average values of some sleep attributes of P1 and P2 in their normal conditions (i.e., beginning of the monitoring) and at the end of the pregnancy. The table also indicates attributes changes (ratio), comparing data at the end of pregnancy to population data and to her own data. As indicated, the ratio of P1 attributes to her own data is higher than the ratio to the population data. On the other hand, the ratio of P2 attributes to her own data is relatively

TABLE 2: P1 and P2 attributes and the ratio of the attributes at the end of pregnancy to her own data and population values.

Attributes	#	Mid of second trimester	End of third trimester	Ratio to popul. data	Ratio to her own data
Sleep fragmentation (times)	P1	0.5	1.53	1.62	3.06
	P2	1.39	2.29	2.43	1.64
WASO (minutes)	P1	15.3	37.32	1.48	2.43
	P2	34.39	75.83	3.02	2.2
Sleep duration (minutes)	P1	389.34	341.25	0.71	0.87
	P2	480.04	456.33	0.95	0.95
Resting heart rate (beats/min)	P1	53.38	59.23	0.96	1.2
	P2	65.61	71.17	1.16	1.08

less.

As shown in Figure 5a, the baseline score is unable to accurately distinguish between P1 and P2. This is because P1's sleep parameters, despite the substantial changes, were close to the population values. In contrast to the baseline method, the sleep changes are clearly visible using the abnormality score obtained from the proposed model (see Figure 5b). This enables the provision of tailored individualized and effective care, where we can identify those who need the care most and optimize resource allocation.

A. LIMITATIONS AND FUTURE DIRECTIONS

The proposed IoT-based system is a proof-of-concept for 1) long-term monitoring of maternal daily sleep 2) effective care for maternal sleep disturbances using personalized decision-making. One of the limitations of this study is that the study sample is small. Other studies investigate the associations between subjective sleep measurements and other pregnancy-related parameters and complications on large study samples. For example, Okun *et al.* [104] conduct a study on 166 pregnant women via self-report questionnaires and indicate that poor sleep quality is correlated with an increased risk of preterm birth. Another study is performed on 457 pregnant women to tackle the association between sleep quality and type of delivery and length of the labor [22]. Unfortunately, we are unable to statistically investigate such associations in our data since our sample size is smaller. Future directions of this study are to perform objective longitudinal studies on a larger population focusing on such correlations.

Another limitation of our monitoring study is linked to the data collection. We were bounded to one wristband that monitored heart rate, step counts, and body movements. Future work will consider multimodal and multisensor data collection and integration with more advanced sensor nodes, enabling the capture of additional health/sleep attributes. For instance, PPG as a non-invasive and convenient technique can play a significant role in such monitoring systems [69]. Finger-based and wrist-based PPG sensors can be leveraged in this regard to continuously acquire different health parameters such as heart rate variability and respiration rate.

Moreover, strap monitors can be employed to record EMG signals for possible abdominal contractions extraction. However, to enhance the feasibility of long-term monitoring, there needs to be a balance between the number of wearables and their continuous use, as a high number of wearable devices could be impractical or inapplicable for sustained long-term monitoring. For instance, in our study, despite using only one wristband for the data collection, we were required to exclude the sleep data of 7 participants out of 20 due to the high volume of missing data. The main reasons were forgetfulness and refusal of wearing the wristband during sleep.

Finally, it is worth noting that the proposed model can be extended to *contextual anomalies* methods, considering the contextual information. These longitudinal studies demand remote and in-home monitoring in which the participants might be involved in different conditions and environments. Therefore, context information including personal lifelogging data, ambient data, and medication reports can improve the accuracy of the personalized decision-making.

VII. CONCLUSION

Maternal sleep quality alters during the pregnancy and postpartum due to the adaptations of the maternal body. Such variations in sleep should be closely monitored as poor sleep quality might lead to various pregnancy complications. Conventional studies are insufficient for this issue as they are limited to restricted data collection approaches. In this paper, we conducted an objective longitudinal study to thoroughly investigate maternal sleep adaptations in pregnancy and postpartum. We introduced an IoT-based system to remotely monitor pregnant women 24/7. Several sleep attributes were extracted to observe changes in maternal sleep patterns. Moreover, we proposed a Bayesian RNN approach to construct a personalized sleep model for each individual using her own data. The sleep model was utilized to deliver an abnormality score, which indicated the degree of maternal sleep quality adaptations. In total, we collected 7 months of data from 20 pregnant women; however, we only included 172.15 ± 33.29 days of valid sleep data per person from 13 pregnant women in our sleep analysis. For each subject, the sleep model was created using the data from the beginning of the monitoring, and the model was tested on the rest of the pregnancy and postpartum data. The obtained scores showed that sleep abnormalities increased during the pregnancy (2.87 times) and after the delivery (5.62 times) in comparison to the mid of the second trimester. This work indicated sleep quality decreased in pregnancy and postpartum with a high confidence level, leveraging fine-grained quantitative measurements and analysis on everyday data of pregnant women.

ACKNOWLEDGMENT

This work was partially supported by the Academy of Finland grants 313448 and 313449 (PREVENT project) and grants 316810 and 316811 (SLIM project) and by the US National Science Foundation (NSF) grant SCC CNS-

1831918 (UNITE project). The authors would like to thank Mr. Arman Anzanpour for designing Figure 1.

REFERENCES

- [1] P. Soma-Pillay et al., "Physiological changes in pregnancy," *Cardiovasc J Afr.*, vol. 27, no. 2, pp. 89–94, 2016.
- [2] K. Lain et al., "Metabolic changes in pregnancy," *Clin Obstet Gynecol.*, vol. 50, no. 4, pp. 938–48, 2007.
- [3] G. Pien and R. Schwab, "Sleep disorders during pregnancy," *sleep*, vol. 27, no. 7, pp. 1405–17, 2004.
- [4] K. Lee et al., "Parity and sleep patterns during and after pregnancy," *Obstetrics & Gynecology*, vol. 95, no. 1, pp. 14–18, 2000.
- [5] C. Hedman et al., "Effects of pregnancy on mothers' sleep," *Sleep Medicine*, vol. 3, no. 1, pp. 37–42, 2002.
- [6] M. Manconi et al., "Restless legs syndrome and pregnancy," *Neurology*, vol. 63, no. 6, pp. 1065–9, 2004.
- [7] L. Hunter et al., "A selective review of maternal sleep characteristics in the postpartum period," *J Obstet Gynecol Neonatal Nurs.*, vol. 38, no. 1, pp. 60–8, 2009.
- [8] F. L. Facco et al., "Sleep disturbances in pregnancy," *Obstetrics & Gynecology*, vol. 115, no. 1, pp. 77–83, 2010.
- [9] M. L. Okun et al., "How disturbed sleep may be a risk factor for adverse pregnancy outcomes a hypothesis," *Obstetrical & gynecological survey*, vol. 64, no. 4, p. 273, 2009.
- [10] S. Reutrakul et al., "Sleep disturbances and their relationship to glucose tolerance in pregnancy," *Diabetes care*, vol. 34, no. 11, pp. 2454–7, 2011.
- [11] L. Palagini et al., "Chronic sleep loss during pregnancy as a determinant of stress: impact on pregnancy outcome," *Sleep Med.*, vol. 15, no. 8, pp. 853–9, 2014.
- [12] J. J. Chang et al., "Sleep deprivation during pregnancy and maternal and fetal outcomes: Is there a relationship?" *Sleep Med Rev.*, vol. 14, no. 2, pp. 107–14, 2010.
- [13] G. Bourjeily et al., "Pregnancy and fetal outcomes of symptoms of sleep-disordered breathing," *European Respiratory Journal*, vol. 36, no. 4, pp. 849–55, 2010.
- [14] G. N. Pires et al., "Sleep impairment during pregnancy: Possible implications on mother–infant relationship," *Medical Hypotheses*, vol. 75, no. 6, pp. 578–82, 2010.
- [15] H. Skouteris et al., "Sleep quality and depression during pregnancy: a prospective study," *J Sleep Res.*, vol. 17, no. 2, pp. 217–20, 2008.
- [16] J. Mindell et al., "Sleep patterns and sleep disturbances across pregnancy," *Sleep Medicine*, vol. 16, no. 4, pp. 483–8, 2015.
- [17] G. Landry et al., "Measuring sleep quality in older adults: a comparison using subjective and objective methods," *Front Aging Neurosci.*, vol. 7, p. 166, 2015.
- [18] D. J. Buysse et al., "The pittsburgh sleep quality index: a new instrument for psychiatric practice and research," *Psychiatry Res.*, vol. 28, no. 2, pp. 193–213, 1989.
- [19] M. Thurtell et al., "The berlin questionnaire screens for obstructive sleep apnea in idiopathic intracranial hypertension," *J Neuroophthalmol*, vol. 31, no. 4, pp. 316–9, 2011.
- [20] B. Gelaye et al., "Poor sleep quality, antepartum depression and suicidal ideation among pregnant women," *Journal of Affective Disorders*, vol. 209, pp. 195–200, 2017.
- [21] G. Li et al., "Relationship between prenatal maternal stress and sleep quality in chinese pregnant women: the mediation effect of resilience," *Sleep Medicine*, vol. 25, pp. 8–12, 2016.
- [22] N. Zafarghandi et al., "The effects of sleep quality and duration in late pregnancy on labor and fetal outcome," *J Matern Fetal Neonatal Med.*, vol. 25, no. 5, pp. 535–7, 2012.
- [23] A. Harvey et al., "The subjective meaning of sleep quality: A comparison of individuals with and without insomnia," *Sleep*, vol. 31, no. 3, pp. 383–93, 2008.
- [24] J. Hartmann et al., "Exploring the construct of subjective sleep quality in patients with insomnia," *J Clin Psychiatry*, vol. 76, no. 6, pp. 768–73, 2015.
- [25] L. Bakken et al., "Subjective sleep quality in relation to objective sleep estimates: comparison, gender differences and changes between the acute phase and the six-month follow-up after stroke," *J Adv Nurs.*, vol. 70, no. 3, pp. 639–50, 2014.
- [26] R. Berry et al., "Polysomnography I," in *Sleep Medicine Pearls*, 3rd ed. Elsevier Health Sciences, 2015, pp. 80–7.
- [27] A. Axelín et al., "Effects of pain management on sleep in preterm infants," *Eur J Pain.*, vol. 14, no. 7, pp. 752–8, 2010.
- [28] S. Ancoli-Israel et al., "The role of actigraphy in the study of sleep and circadian rhythms," *Sleep*, vol. 26, no. 3, pp. 342–92, 2003.
- [29] K. Lee and C. Gay, "Sleep in late pregnancy predicts length of labor and type of delivery," *Am J Obstet Gynecol.*, vol. 191, no. 6, pp. 2041–6, 2004.
- [30] S. Coo Calcagni et al., "The relationship between sleep and mood in first-time and experienced mothers," *Behav Sleep Med.*, vol. 10, no. 3, pp. 167–79, 2012.
- [31] S. Tsai et al., "Cross-sectional and longitudinal associations between sleep and health-related quality of life in pregnant women: a prospective observational study," *Int J Nurs Stud.*, vol. 56, pp. 45–53, 2016.
- [32] A. Haney et al., "Sleep disturbance and cardiometabolic risk factors in early pregnancy: a preliminary study," *Sleep Medicine*, vol. 15, no. 4, pp. 444–50, 2014.
- [33] L. Atzori et al., "The internet of things: A survey," *Computer Networks*, vol. 54, no. 15, pp. 2787–805, 2010.
- [34] J. Gubbi et al., "Internet of things (iot): A vision, architectural elements, and future directions," *Future Generation Computer Systems*, vol. 29, no. 7, pp. 1645–60, 2013.
- [35] M. Chen et al., "Big data: A survey," *Mobile Networks and Applications*, vol. 19, no. 2, pp. 171–209, 2014.
- [36] J. Catov and C. Margerison-Zilko, "Pregnancy as a window to future health: short term costs and consequences," *Am J Obstet Gynecol.*, vol. 215, no. 4, pp. 406–7, 2017.
- [37] UNICEF, "Antenatal Care - UNICEF DATA," Retrieved on July 2019, <https://data.unicef.org/topic/maternal-health/antenatal-care/>.
- [38] Y. Yang et al., "Determinants of sleep quality among pregnant women in china: a cross-sectional survey," *J Matern Fetal Neonatal Med.*, vol. 31, no. 22, pp. 2980–5, 2017.
- [39] B. Bei et al., "Sleep and mood during pregnancy and the postpartum period," *Sleep Medicine Clinics*, vol. 10, no. 1, pp. 25–33, 2015.
- [40] B. Izci Balserak and K. Lee, "Sleep disturbances and sleep-related disorders in pregnancy," in *Principles and practice of sleep medicine*. Elsevier, 2010, pp. 1572–86.
- [41] V. Ibáñez et al., "A survey on sleep assessment methods," *PeerJ*, vol. 6, p. e4849, 2018.
- [42] H. Zhou et al., "Poor sleep quality of third trimester exacerbates the risk of experiencing postnatal depression," *Psychol Health Med.*, vol. 17, pp. 1–10, 2018.
- [43] E. Paavonen et al., "Maternal and paternal sleep during pregnancy in the child-sleep birth cohort," *Sleep Medicine*, vol. 29, pp. 47–56, 2017.
- [44] S. Herring et al., "Do pregnant women accurately report sleep time?" *Sleep and Breathing*, vol. 17, no. 4, pp. 1323–7, 2013.
- [45] B. Vaughn and P. Giallanza, "Technical review of polysomnography," *Chest*, vol. 134, no. 6, pp. 1310–9, 2008.
- [46] S. Herring et al., "Objectively measured sleep duration and hyperglycemia in pregnancy," *Sleep Medicine*, vol. 15, no. 1, pp. 51–5, 2014.
- [47] F. Lin et al., "Sleepsense: A noncontact and cost-effective sleep monitoring system," *IEEE Trans Biomed Circuits Syst.*, vol. 11, no. 1, pp. 189–202, 2017.
- [48] M. Tenhunen et al., "Emfit movement sensor in evaluating nocturnal breathing," *Respir Physiol Neurobiol.*, vol. 187, no. 2, pp. 183–9, 2013.
- [49] S. Saeb et al., "Scalable passive sleep monitoring using mobile phones: Opportunities and obstacles," *J Med Internet Res.*, vol. 19, no. 4, p. e118, 2017.
- [50] Z. Liu et al., "Characterization of maternal body movement during sleep before and after parturition," in *Conf Proc IEEE Eng Med Biol Soc.*, 2011, pp. 1503–6.
- [51] V. Chandola et al., "Anomaly detection: A survey," *ACM Comput. Surv.*, vol. 41, no. 3, pp. 15:1–15:58, 2009.
- [52] M. Ahmed et al., "A survey of network anomaly detection techniques," *J. of Network and Computer Applications*, vol. 60, pp. 19–31, 2016.
- [53] S. Han and S. Cho, "Evolutionary neural networks for anomaly detection based on the behavior of a program," *IEEE Trans. on Systems, Man, and Cybernetics*, vol. 36, no. 3, pp. 559–70, 2005.
- [54] W. Chen et al., "Application of SVM and ANN for intrusion detection," *Computers & Operations Research*, vol. 32, no. 10, pp. 2617–34, 2005.
- [55] W. Fan et al., "Using artificial anomalies to detect unknown and known network intrusions," in *IEEE Int. Conf. on Data Mining*, 2001, pp. 123–130.

- [56] T. Fawcett and F. Provost, "Activity monitoring: Noticing interesting changes in behavior," in *Fifth ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining*, ser. KDD '99, 1999, pp. 53–62.
- [57] L. Manevitz and M. Yousef, "One-class SVMs for document classification," *J. Mach. Learn. Res.*, vol. 2, pp. 139–154, 2002.
- [58] R. Chalapathy, A. K. Menon, and S. Chawla, "Anomaly detection using one-class neural networks," *CoRR*, vol. abs/1802.06360, 2018.
- [59] S. Hawkins et al., "Outlier detection using replicator neural networks," in *Int. Conf. on Data Warehousing and Knowledge Discovery*, 2002, pp. 170–80.
- [60] H. Dau et al., "Anomaly detection using replicator neural networks trained on examples of one class," in *Asia-Pacific Conf. on Simulated Evolution and Learning*, 2014, pp. 311–22.
- [61] S. Basu et al., "A probabilistic framework for semi-supervised clustering," in *ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining*, 2004, pp. 59–68.
- [62] J. Lin et al., "Approximations to magic: Finding unusual medical time series," in *IEEE Symposium on CBMS*, 2005, pp. 329–34.
- [63] L. Cadmus-Bertram, "Using fitness trackers in clinical research: What nurse practitioners need to know," *J Nurse Pract.*, vol. 13, no. 1, pp. 34–40, 2017.
- [64] M. de Zambotti et al., "The sleep of the ring: comparison of the ōura sleep tracker against polysomnography," *Behavioral sleep medicine*, vol. 17, no. 2, pp. 124–136, 2019.
- [65] S. Kang et al., "Validity of a commercial wearable sleep tracker in adult insomnia disorder patients and good sleepers," *J. of psychosomatic research*, vol. 97, pp. 38–44, 2017.
- [66] M. de Zambotti et al., "Validation of sleep-tracking technology compared with polysomnography in adolescents," *Sleep*, vol. 38, no. 9, pp. 1461–1468, 2015.
- [67] Garmin, "Vivosmart HR | Activity Tracker," Retrieved on July 2019, <https://buy.garmin.com/en-US/US/p/531166>.
- [68] K. Grym et al., "Feasibility of smart wristbands for continuous monitoring during pregnancy and one month after birth," *BMC Pregnancy and Childbirth*, vol. 19, no. 34, 2019.
- [69] J. Allen, "Photoplethysmography and its application in clinical physiological measurement," *Physiol Meas.*, vol. 28, no. 3, pp. 1–39, 2007.
- [70] Linode, "Ssd cloud hosting & linux servers," Retrieved on July 2019, <https://www.linode.com/>.
- [71] A. Sadeh, "III. Sleep assessment methods," *Monographs of the Society for Research in Child Development*, vol. 80, no. 1, pp. 33–48, 2015.
- [72] L. Tikotzky and E. Volkovich, "Infant nocturnal wakefulness: a longitudinal study comparing three sleep assessment methods," *Sleep*, vol. 42, no. 1, p. zsy191, 2018.
- [73] V. Natale et al., "Actigraphy in the assessment of insomnia: a quantitative approach," *Sleep*, vol. 32, no. 6, pp. 767–71, 2009.
- [74] I. Fonareva et al., "Assessing sleep architecture in dementia caregivers at home using an ambulatory polysomnographic system," *Journal of Geriatric Psychiatry and Neurology*, vol. 24, no. 1, pp. 50–9, 2011.
- [75] A. D. Krystal and J. D. Edinger, "Measuring sleep quality," *Sleep Medicine*, vol. 9, no. 1, pp. s10–7, 2008.
- [76] C. Kline, "Sleep duration," in Gellman M.D., Turner J.R. (eds) *Encyclopedia of Behavioral Medicine*. Springer, 2013, pp. 1808–10.
- [77] M. Jike et al., "Long sleep duration and health outcomes: A systematic review, meta-analysis and meta-regression," *Sleep Medicine Reviews*, vol. 39, pp. 25–36, 2018.
- [78] A. Vgontzas et al., "Insomnia with objective short sleep duration is associated with a high risk for hypertension," *Sleep.*, vol. 32, no. 4, pp. 491–7, 2009.
- [79] X. Zhou et al., *Applied missing data analysis in the health sciences*. John Wiley & Sons, Inc., 2014.
- [80] B. Balserak and K. Lee, "Sleep and sleep disorders associated with pregnancy," in *Principles and Practice of Sleep Medicine*. Elsevier, 2017, pp. 1525–39.
- [81] E. Mezick, "Sleep fragmentation," in Gellman M.D., Turner J.R. (eds) *Encyclopedia of Behavioral Medicine*. Springer, 2013, pp. 1810–1.
- [82] C. Kline, "Sleep quality," in Gellman M.D., Turner J.R. (eds) *Encyclopedia of Behavioral Medicine*. Springer, 2013, pp. 1811–3.
- [83] S. Insana, "Non-REM sleep," in Gellman M.D., Turner J.R. (eds) *Encyclopedia of Behavioral Medicine*. Springer, 2013, pp. 1342–3.
- [84] M. Cooney et al., "Elevated resting heart rate is an independent risk factor for cardiovascular disease in healthy men and women," *American Heart Journal*, vol. 159, no. 4, pp. 612–9.e3, 2010.
- [85] M. Bohm et al., "Resting heart rate: Risk indicator and emerging risk factor in cardiovascular disease," *The American Journal of Medicine*, vol. 128, no. 3, pp. 219–28, 2015.
- [86] C. Zhang et al., "A deep neural network for unsupervised anomaly detection and diagnosis in multivariate time series data," *CoRR*, vol. abs/1811.08055, 2018.
- [87] C. Zhou and R. Paffenroth, "Anomaly detection with robust deep autoencoders," in *ACM Int. Conf. on Knowledge Discovery and Data Mining*, 2017, pp. 665–674.
- [88] Y. Gal, "Uncertainty in deep learning," Ph.D. dissertation, University of Cambridge, 2016.
- [89] D. MacKay, "A practical bayesian framework for backpropagation networks," *Neural Comput.*, vol. 4, no. 3, pp. 448–72, 1992.
- [90] R. Neal, *Bayesian Learning for Neural Networks*. Springer Science & Business Media, 2012.
- [91] T. Nguyen and G. Armitage, "A survey of techniques for internet traffic classification using machine learning," *IEEE Communications Surveys & Tutorials*, vol. 10, pp. 56–76, 2008.
- [92] D. Blei et al., "Variational Inference: A Review for Statisticians," *J. of the American Statistical Association*, vol. 112, no. 518, pp. 859–77, 2017.
- [93] M. Jordan et al., "An introduction to variational methods for graphical models," *Machine Learning*, vol. 37, no. 2, pp. 183–233, 1999.
- [94] S. Kullback and R. Leibler, "On information and sufficiency," *Ann. Math. Statist.*, vol. 22, no. 1, pp. 79–86, 1951.
- [95] A. Kucukelbir et al., "Automatic Differentiation Variational Inference," *The J. of Machine Learning Research*, vol. 18, no. 1, pp. 430–74, 2017.
- [96] Lasagne Contributors, "Lasagne 0.2.dev1 documentation," Retrieved on July 2019, <https://lasagne.readthedocs.io/en/latest/>.
- [97] PyMC Development Team, "PyMC3 documentation," Retrieved on July 2019, <https://docs.pymc.io/>.
- [98] O. Hughes et al., "The significance of breastfeeding on sleep patterns during the first 48 hours postpartum for first time mothers," *J Obstet Gynaecol*, vol. 38, no. 3, pp. 316–20, 2018.
- [99] J. Doering et al., "Sleep quality and quantity in low-income postpartum women," *MCN Am J Maternal Child Nurs*, vol. 42, no. 3, pp. 166–72, 2017.
- [100] R. Roman-Galvez et al., "Factors associated with insomnia in pregnancy: A prospective cohort study," *Eur J Obstet Gynecol Reprod Biol.*, vol. 221, pp. 70–75, 2018.
- [101] P. Polo-Kantola et al., "Sleep quality during pregnancy: association with depressive and anxiety symptoms," *Acta Obstet Gynecol Scand*, vol. 96, no. 2, pp. 198–206, 2017.
- [102] L. Wolfe and T. Weissgerber, "Clinical physiology of exercise in pregnancy: a literature review," *J Obstet Gynaecol Can*, vol. 25, no. 6, pp. 473–83, 2003.
- [103] I. Sedov et al., "Sleep quality during pregnancy: A meta-analysis," *Sleep Med Rev.*, vol. 38, pp. 168–76, 2018.
- [104] M. Okun et al., "Poor sleep quality is associated with preterm birth," *Sleep*, vol. 34, no. 11, pp. 1493–8, 2011.

...